ELSEVIER Linear Algebra and its Applications 285 (1998) 229–255

LINEAR ALGEBRA AND ITS APPLICATIONS

Solution of Toeplitz normal equations by sine transform based preconditioning

Fabio Di Benedetto [1]

Dipartimento di Matematica, Università di Genova, via Dodecaneso 35, I-16146 Genova, Italy

Received 3 November 1996; accepted 29 June 1998

Communicated by G. Heinig

Abstract

The normal equations constructed by a Toeplitz matrix are studied, in order to find a suitable preconditioner related to the discrete sine transform. New results are given about the structure of the product of two Toeplitz matrices, which allow the CGN method to achieve a superlinear rate of convergence. This preconditioner outperforms the circulant one for the iterative solution of Toeplitz least-squares problems; such strategy can also be applied to nonsymmetric linear systems. A block generalization is discussed. © 1998 Elsevier Science Inc. All rights reserved.

AMS classification: 65F10; 65F15

Keywords: Toeplitz matrix; Least-squares; Normal equations; Preconditioning; Sine transform

1. Introduction

Let $C_{m,n} \in \mathbb{R}^{m\times n}$ be a *Toeplitz matrix*, that is its (i,j) element is a function c_{i-j} of the difference of indices. When the elements $\{c_k\}$ on each diagonal come from the formal Laurent expansion of a function $r(z) = \sum_{k=-\infty}^{+\infty} c_k z^k$ on the unit circle, we use the notation $C_{m,n} = T_{m,n}(r)$ (or $T_n(r)$ if $m = n$) and r is called the *generating function* of $C_{m,n}$.

[1] E-mail: dibenede@dima.unige.it.

PII: S0024-3795(98)10115-5

A classical assumption considers r belonging to the *Wiener class*, for which the condition $\sum_{k=-\infty}^{+\infty} |c_k| < +\infty$ holds. We characterize in this way a proper subset of continuous functions.

The singular values of $C_{m,n}$ are distributed (in the sense specified by Tyrtyshnikov [32]) as the values of the function $r(z)$, as shown in [28,32]. If $r(z)$ vanishes at some point of the unit circle then the singular values tend to accumulate at the origin, and we therefore say that the matrices $\{C_{m,n}\}$ are *asymptotically ill-conditioned.*

In many applications, especially those involving image restoration [7,27], it is required to solve a least-squares problem like

$$\min_{\mathbf{x}} \|C_{m,n}\mathbf{x} - \mathbf{y}\|_2, \tag{1}$$

where $C_{m,n}$ is a $m \times n$ (block) Toeplitz matrix, often ill-conditioned.

There is an extensive literature about numerical methods for treating the case where $m = n$ and $T_n(r)$ is nonsingular

$$T_n(r)\mathbf{x} = \mathbf{y}. \tag{2}$$

Recently, particular attention has been paid for preconditioning techniques (see the survey [11]) allowing one to compute the solution in a parallel logarithmic cost per iteration; on the other hand, "superfast" direct methods [5,1] are intrinsically sequential and could be unstable in presence of nonsymmetry.

The first idea of choosing the preconditioner P_n for $T_n(r)$ in the circulant class [13,14,24,23,6], for which P_n^{-1} is computed by the discrete Fourier transform, has been generalized in order to solve a least-squares problem. More precisely, a "circulant approximation" $P_{m,n}$ of $C_{m,n}$ can be used for this purpose, by proving that $P_{m,n}^{\mathrm{T}} P_{m,n}$ is a good preconditioner for the *normal equations*

$$C_{m,n}^{\mathrm{T}} C_{m,n}\mathbf{x} = C_{m,n}^{\mathrm{T}}\mathbf{y}, \tag{3}$$

such theoretical results hold under the (strong) assumption that $|r(z)| > 0$ [8,7].

A different approach tries to find a circulant approximation to the coefficient matrix in Eq. (3), but does not remove the positivity assumption [9].

The $\mathscr{T}$ class [3,18,17], related to the discrete sine transform, is another family of efficient preconditioners for Toeplitz matrices. With this choice it is possible to solve symmetric, ill-conditioned Toeplitz systems in a number of iterations not depending on the dimension [15,19], and this is not achieved by the use of a circulant preconditioner.

Unfortunately, the $\mathscr{T}$ class is intrinsically symmetric and it is hopeless to look for a $\mathscr{T}$ approximation of the matrix $C_{m,n}$ appearing in Eq. (1), in order to improve the performance of the circulant approach.

In this paper we overcome this problem by finding directly a $\mathscr{T}$ approximation P to $R_n = C_{m,n}^{\mathrm{T}} C_{m,n}$, based on the key result that the product of two Toeplitz matrices is no longer Toeplitz, but it preserves this structure after slight corrections.

In particular, for a matrix of the form $T^{\mathrm{T}}T$ (where T is Toeplitz) we find a closely related Toeplitz matrix T_0: if T is banded, then T_0 differs from $T^{\mathrm{T}}T$ by a low-rank update; if T is full, then a further term whose norm is neglectable is added.

We finally prove that choosing P as the standard approximation of T_0 in $\mathscr{T}$ leads to a clustering behaviour of the eigenvalues of $P^{-1}(T^{\mathrm{T}}T)$ around 1, which ensures a fast convergence of the Conjugate Gradient method applied to the Normal equation (CGN [21]).

The above result not only holds under the assumption $|r(z)| > 0$, but also for every banded Toeplitz matrix, though asymptotically ill-conditioned: in this case the circulant preconditioner fails to achieve a superlinear rate of convergence. The same approach can be used also in the case $m = n$, in order to solve the nonsymmetric system (2) by an iterative method; in fact, it is well-known that the CGN method is effective when a good preconditioner of the coefficient matrix in Eq. (3) is at our disposal.

In Section 2 we recall the definition and the main properties of the $\mathscr{T}$ class. Section 3 contains our new results concerning the structure of a product of two Toeplitz matrices, for which the preconditioner is constructed and analyzed; the banded case is considered first, the argument is extended later to full matrices. In Section 4 some numerical examples are reported, while in Section 5 we give an outline of the generalization of such results to block Toeplitz matrices.

Notations. We define $\mathscr{B}_{M,N,p}$ as the set of $np \times np$ matrices having the following pattern

$$E \in \mathscr{B}_{M,N,p} \iff E = \begin{pmatrix} E^{(NW)} & & \\ & 0 & \\ & & E^{(SE)} \end{pmatrix},$$

in $E^{(NW)} \in \mathbb{R}^{Mp \times Mp}$ and $E^{(SE)} \in \mathbb{R}^{Np \times Np}$ are concentrated the only nonzero elements of E, whence $\mathrm{rk}(E) \leqslant (M+N)p$.

The "dual" family $\mathscr{B}_{K,L,n}^{\Pi}$ consists of block matrices $E \in \mathbb{R}^{np \times np}$ partitioned as follows:

$$E = \begin{pmatrix} E_{1,1} & \cdots & E_{1,n} \\ \vdots & & \vdots \\ E_{n,1} & \cdots & E_{n,n} \end{pmatrix},$$

where each $E_{i,j} \in \mathbb{R}^{p \times p}$ belongs to $\mathscr{B}_{K,L,1}$. In this case, $\mathrm{rk}(E)$ does not exceed $(K+L)n$.

2. The algebra $\mathcal{T}$

We will recall in this section the definition and the algebraic properties of the class $\mathcal{T}$, that has been introduced the first time in [2] and generalized later in [26] to the block case.

Let $E_n = \left(\sqrt{(2/n+1)}\sin(\pi ij/(n+1))\right)_{i,j=1}^{n}$ be the $n \times n$ matrix associated to the discrete sine transform; E_n is symmetric and orthogonal. The class $\mathcal{T}_n$ is defined as

$$\mathcal{T}_n = \{P \in \mathbb{R}^{n\times n} : E_n P E_n \text{ is diagonal}\},$$

it is an algebra of symmetric matrices, containing all the Toeplitz tridiagonal matrices as a particular case.

The computation of the eigenvalues of a matrix $P \in \mathcal{T}_n$ from its first column can be performed by a fast sine transform at a cost of $O(n \log n)$ operations or $O(\log n)$ parallel steps with $O(n)$ processors: this can be used in order to solve a linear system associated to P at the same cost.

It is possible to relate a $\mathcal{T}_n$ matrix to a given real-valued Wiener function $f(z)$, through a standard correction of the Toeplitz matrix $T_n(f)$. If

$$\hat{f}(x) = f(e^{ix}) = \sum_{j=-\infty}^{+\infty} c_j e^{ijx} \quad (c_j = c_{-j})$$

is the Laurent expansion of f on the unit circle S^1, define

$$\tau_n(f) = \begin{pmatrix} c_0 & \dots & c_{n-1} \\ \vdots & \ddots & \vdots \\ c_{n-1} & \dots & c_0 \end{pmatrix} - \begin{pmatrix} c_2 & \dots & c_{n-1} & 0 & 0 \\ \vdots & \iddots & & & 0 \\ c_{n-1} & & 0 & & c_{n-1} \\ 0 & & & \iddots & \vdots \\ 0 & 0 & c_{n-1} & \dots & c_2 \end{pmatrix}, \tag{4}$$

observe that the first term is exactly $T_n(f)$, while the second one has the *Hankel structure* (i.e. each entry depends on the sum of the subscripts) and it is persymmetric (i.e. symmetric with respect to the secondary diagonal). In this case, the eigenvalues of $\tau_n(f)$ are given by the values $\{\hat{f}_n(\pi j/(n+1))\}_{j=1,\dots,n}$, where $\hat{f}_n$ is the nth partial Fourier sum of $\hat{f}$. In [3,15] it is studied the behaviour of $\tau_n(f)$ in the preconditioning of the conjugate gradient method applied to $T_n(f)$; the same results have been independently rediscovered in the more recent papers [4,12].

Similarly, we may define the block extension $\mathcal{T}_{n,p}$: if $E_{n,p}$ is the Kronecker product $E_n \otimes E_p$, involving the discrete sine transform in two dimensions, the new definition is

$$\mathcal{T}_{n,p} = \{P \in \mathbb{R}^{np\times np} : E_{n,p} P E_{n,p} \text{ is diagonal}\}.$$

Now, the solution of a linear system in $\mathscr{T}_{n,p}$ can be solved with $O(np \log(np))$ operations or $O(\log n + \log p)$ parallel steps and $O(np)$ processors.

Every matrix $P \in \mathscr{T}_{n,p}$ is *quadrantally symmetric*; that is, if P is partitioned as

$$P = \begin{pmatrix} p_{1,1} & \cdots & p_{1,n} \\ \vdots & & \vdots \\ p_{n,1} & \cdots & p_{n,n} \end{pmatrix} \quad \text{with } p_{i,j} \in \mathbb{R}^{p\times p},$$

then the conditions $p_{i,j} = p_{j,i}$ and $p_{i,j} = p_{i,j}^{\mathrm{T}}$ hold.

As in the scalar case, from the bivariate Laurent expansion

$$\hat{\phi}(x,y) = \phi(e^{ix}, e^{iy}) = \sum_{j,k=-\infty}^{+\infty} c_{j,k} e^{ijx} e^{iky}$$

on the Cartesian product $S^1 \times S^1$ we can define

$$\tau_{n,p}[\phi] = \begin{pmatrix} c_0 & \cdots & c_{n-1} \\ \vdots & \ddots & \vdots \\ c_{n-1} & \cdots & c_0 \end{pmatrix} - \begin{pmatrix} c_2 & \cdots & c_{n-1} & 0 & 0 \\ \vdots & \iddots & & & 0 \\ c_{n-1} & & 0 & & c_{n-1} \\ 0 & & & \iddots & \vdots \\ 0 & 0 & c_{n-1} & \cdots & c_2 \end{pmatrix},$$

where

$$c_j = \begin{pmatrix} c_{j,0} & \cdots & c_{j,p-1} \\ \vdots & \ddots & \vdots \\ c_{j,p-1} & \cdots & c_{j,0} \end{pmatrix} - \begin{pmatrix} c_{j,2} & \cdots & c_{j,p-1} & 0 & 0 \\ \vdots & \iddots & & & 0 \\ c_{j,p-1} & & 0 & & c_{j,p-1} \\ 0 & & & \iddots & \vdots \\ 0 & 0 & c_{j,p-1} & \cdots & c_{j,2} \end{pmatrix}.$$

In [17] it is shown that the eigenvalues of $\tau_{n,p}[\phi]$ are related to the ordinates $\{\hat{\phi}(\pi j/(n+1), \pi k/(p+1))\}_{j,k}$; the preconditioning properties of $\mathscr{T}_{n,p}$ are also studied in the same work.

The classes $\mathscr{T}_n$ and $\mathscr{T}_{n,p}$ are well-suited for preconditioning Toeplitz matrices that are (quadrantally) symmetric and positive definite; on the other hand, symmetry is a strong restriction for their use in more general settings. The results of the next section will enlarge the range of applicability of the algebra $\mathscr{T}$.

3. Main results

3.1. The banded case

Theorem 3.1. *Let $a(z)$ and $b(z)$ be Laurent polynomials such that:*

$$a(z) = a_{-N}z^{-N} + \cdots + a_0 + \cdots + a_M z^M,$$
$$b(z) = b_{-M}z^{-M} + \cdots + b_0 + \cdots + b_N z^N,$$

then the matrices $T_n(a \cdot b)$ and $T_{n,m}(a) \cdot T_{m,n}(b)$ agree except their $M \times M$ north–west and $N \times N$ south–east corners; that is, their difference is a $\mathscr{B}_{M,N,1}$ matrix.

Proof. Let $k = M + N$; define A, B as the band upper triangular Toeplitz matrices of order $n \times (m + k)$

$$A = \begin{pmatrix} a_M & \cdots & a_0 & \cdots & a_{-N} & & 0 \\ & \ddots & & \ddots & & \ddots & \\ 0 & & a_M & \cdots & a_0 & \cdots & a_{-N} \end{pmatrix},$$

$$B = \begin{pmatrix} b_{-M} & \cdots & b_0 & \cdots & b_N & & 0 \\ & \ddots & & \ddots & & \ddots & \\ 0 & & b_{-M} & \cdots & b_0 & \cdots & b_N \end{pmatrix}.$$

The matrices A and B can be partitioned as

$$A = \begin{pmatrix} U_A & & 0 \\ 0 & T_{n,m}(a) & 0 \\ 0 & & L_A \end{pmatrix}, \quad B = \begin{pmatrix} U_B & & 0 \\ 0 & T_{n,m}(b)^{\mathrm{T}} & 0 \\ 0 & & L_B \end{pmatrix},$$

where U_A, U_B are $M \times M$ upper triangular, L_A, L_B are $N \times N$ lower triangular Toeplitz matrices (provided that $N \leqslant m - n$, the latter blocks vanish).

Then we are able to express the matrix product $A \cdot B^{\mathrm{T}}$ as

$$AB^{\mathrm{T}} = \begin{pmatrix} U_A U_B^{\mathrm{T}} & 0 \\ 0 & 0 \end{pmatrix} + T_{n,m}(a)T_{m,n}(b) + \begin{pmatrix} 0 & 0 \\ 0 & L_A L_B^{\mathrm{T}} \end{pmatrix}, \tag{5}$$

but its (i, j) element is also directly given, for $0 \leqslant i, j \leqslant n - 1$, by

$$\sum_{l=0}^{m+k-1} a_{i-l+M} b_{l-j-M} = \sum_{r=-N}^{M} a_r b_{i-j-r} = c_{i-j},$$

which is exactly the coefficient of z^{i-j} in $a(z)b(z)$; this implies

$$AB^{\mathrm{T}} = T_n(ab). \tag{6}$$

Comparing Eqs. (5) and (6) yields the thesis. □

If $N = M$, $b(z) = a(z^{-1})$ and $T_{m,n}(b)$ has full rank, then the square symmetric matrix $T_{m,n}(b)^{\mathrm{T}}T_{m,n}(b) = T_{n,m}(a)T_{m,n}(b)$ is positive definite; by Theorem 3.1, such matrix is close to the spd Toeplitz matrix $T_n(ab)$. So we can apply the correction $\tau_n(ab)$ of $T_n(ab)$ in $\mathscr{T}$ for preconditioning the matrix $T_{m,n}(b)^{\mathrm{T}}T_{m,n}(b)$. In other words, we propose to apply the CG method to the normal equations (3) by using $\tau_n(ab)$ as a preconditioner. When $n = m$, this reduces to using a preconditioned CGN method for solving the nonsymmetric linear system (2). The effectiveness of this strategy is ensured by the following result:

Theorem 3.2. *Let* $R_n = T_{m,n}(b)^{\mathrm{T}}T_{m,n}(b)$, $P_n = \tau_n(ab)$ *as above. Then*

$$P_n^{-1}R_n = I_n + E_n \quad \textit{with } \mathrm{rk}(E_n) \leqslant 4 \cdot \max(d^+, d^-) - 2,$$

where d^+ *and* d^- *are the degrees of the Laurent polynomial* $b(z)$. *Hence, the CG method applied to* $R_n\mathbf{x} = T_{m,n}(b)^{\mathrm{T}}\mathbf{y}$ *computes the solution in exact arithmetic in* $4 \cdot \max(d^+, d^-) - 2$ *iterations.*

Proof. If $\Delta_n = R_n - P_n$, we have $P_n^{-1}R_n = I_n + P_n^{-1}\Delta_n$, so that it suffices to prove that $\mathrm{rk}(\Delta_n) \leqslant 4 \cdot \max(d^+, d^-) - 2$. Write

$$\Delta_n = (T_{m,n}(b)^{\mathrm{T}}T_{m,n}(b) - T_n(ab)) + (T_n(ab) - \tau_n(ab)) \tag{7}$$

and observe that the first difference of Eq. (7) vanishes except in the north-west and south-east corners of maximal size $N = \max(d^+, d^-)$, as we have seen in Theorem 3.1. By Eq. (4), the second difference equals a persymmetric Hankel matrix whose nonzero elements are concentrated in the leading and trailing principal submatrices of order $2N - 1$, that is the bandwidth of $T_n(ab)$. Then Δ_n has a pattern quite similar to the latter, and its rank cannot exceed $4N - 2$.

Now it suffices to invoke a well-known convergence result on the conjugate gradient method [20] to prove the thesis. □

Remark. No assumption is made in Theorem 3.2 about the condition number of $T_{m,n}(b)$. The result is true even though the generating function $b(z)$ vanishes somewhere on the unit circle.

3.2. The full case

The preconditioning technique described so far is effective for full Toeplitz matrices too: if $T_{m,n}(f)$ is the rectangular Toeplitz matrix generated by the function f which is not a polynomial, we still propose to apply the CG method to $R_n = T_{m,n}(f)^{\mathrm{T}}T_{m,n}(f)$ by using the matrix $P_n = \tau_n(gf) \in \mathscr{T}$ as a precondi-

tioner, with $g(z) = f(z^{-1})$. The main difference with respect to the banded case is that the eigenvalues of $P_n^{-1}T_n$ no longer equal 1 exactly, but they cluster around unity. In order to prove this, we need to estimate the norm of a Toeplitz matrix in relation to its generating function.

Lemma 3.1. *Let $r(z)$ be a complex-valued function on the unit circle S^1 belonging to L^∞; then the inequality*

$$\|T_{m,n}(r)\|_2 \leqslant \|r\|_\infty$$

holds, where $\|r\|_\infty$ denotes its supremum norm on S^1.

Proof. It is a straightforward generalization to the rectangular case of the Corollary of Lemma 4.2 in [28]. □

We are ready to compare the matrix R_n to the Toeplitz matrix $T_n(gf)$; the following result deals with the more general case where g and f are not related.

Theorem 3.3. *If $f(z)$ and $g(z)$ are Wiener functions, for every $\epsilon > 0$ there exist N, M such that*

$$T_{n,m}(g)T_{m,n}(f) - T_n(gf) = E_n' + E_n'' \tag{8}$$

for all n sufficiently large, with $\|E_n''\|_2 \leqslant \epsilon$ and $E_n' \in \mathscr{B}_{M,N,1}$.

Proof. Since $f(z)$ and $g(z)$ belong to the Wiener class, for all $\epsilon > 0$ we can define polynomials $a(z)$ and $b(z)$ coming from a suitable truncation of the Laurent series of $f(z)$ and $g(z)$, such that

$$\|a\|_\infty \leqslant 2\|f\|_\infty,\ \|b\|_\infty \leqslant 2\|g\|_\infty,\ \|f-a\|_\infty < \delta,\ \|g-b\|_\infty < \delta, \tag{9}$$

with $\delta = \epsilon/3(\|f\|_\infty + \|g\|_\infty)$. The product $g \cdot f$ is also well approximated by $b \cdot a$, since

$$\begin{aligned}\|gf - ba\|_\infty = \|gf - ga + ga - ba\|_\infty &\leqslant \|g\|_\infty\|f-a\|_\infty + \|a\|_\infty\|g-b\|_\infty \\ &< \delta(\|g\|_\infty + 2\|f\|_\infty).\end{aligned} \tag{10}$$

Split now the difference in Eq. (8) as

$$T_{n,m}(g)T_{m,n}(f) - T_n(gf) = E_n' + F_n + G_n$$

with:

$$\begin{aligned}E_n' &= T_{n,m}(b)T_{m,n}(a) - T_n(ba),\\ F_n &= T_n(ba) - T_n(gf),\\ G_n &= T_{n,m}(g)T_{m,n}(f) - T_{n,m}(b)T_{m,n}(a).\end{aligned}$$

If the degrees N and M of $a(z)$ and $b(z)$ are chosen according to the assumptions of Theorem 3.1, we immediately deduce that the matrix E_n' has the desired pattern. Since F_n equals the Toeplitz matrix generated by $b(z)a(z) - g(z)f(z)$, by Lemma 3.1 and Eq. (10) we have

$$\|F_n\|_2 < \delta(\|g\|_\infty + 2\|f\|_\infty).$$

Finally, we can write

$$\begin{aligned} G_n &= T_{n,m}(g)T_{m,n}(f) - T_{n,m}(b)T_{m,n}(f) + T_{n,m}(b)T_{m,n}(f) - T_{n,m}(b)T_{m,n}(a) \\ &= T_{n,m}(g-b)T_{m,n}(f) + T_{n,m}(b)T_{m,n}(f-a), \end{aligned}$$

whence

$$\begin{aligned} \|G_n\|_2 &\leqslant \|T_{n,m}(g-b)\|_2\|T_{m,n}(f)\|_2 + \|T_{n,m}(b)\|_2\|T_{m,n}(f-a)\|_2 \\ &\leqslant \|g-b\|_\infty\|f\|_\infty + \|b\|_\infty\|f-a\|_\infty < \delta(\|f\|_\infty + 2\|g\|_\infty) \end{aligned}$$

in view of Eq. (9). We can so derive that the matrix $E_n'' = F_n + G_n$ satisfies

$$\|E_n''\|_2 \leqslant \|F_n\|_2 + \|G_n\|_2 < 3\delta(\|f\|_\infty + \|g\|_\infty) = \epsilon,$$

that completes the proof of (8). □

This result is new but has interesting connections with the existing literature. A recent theoretical result of Tyrtyshnikov [32] claims that finding E_n' and E_n'' such that

$$\mathrm{rk}(E_n') = o(n) \quad \text{and} \quad \|E_n''\|_F^2 = o(n) \tag{11}$$

is sufficient to prove that the singular values of $T_{n,m}(g)T_{m,n}(f)$ and $T_n(gf)$ are "equally distributed" in the sense of Weyl (see [32] for more details).

The same author provides in [31] a different proof for this equal distribution, under the more general assumption of f and g belonging to L^∞. On the other hand, our relation appearing in the statement of the previous theorem is much stronger than Eq. (11), and this allows us to obtain a superlinear convergence rate for our preconditioning technique, as shown in Theorem 3.4 below.

Moreover, Theorem 3.3 was already known in the particular case $n = m$, $g = 1/f$ [10], which led to the proposal of preconditioning the Hermitian Toeplitz matrices $T_n(f)$ by $T_n(1/f)^{-1}$. Instead, we want to set here $g(z) = \overline{f(z)}$ in order to prove the clustering of the preconditioned spectrum for our CGN method.

Theorem 3.4. *Let $f(z)$ be a Wiener function with no root on the unit circle; setting $R_n = T_{m,n}(f)^{\mathrm{T}}T_{m,n}(f)$ and $P_n = \tau_n(|f|^2)$, then the eigenvalues of $P_n^{-1}R_n$ satisfy the clustering property*

$$\forall \epsilon\ \exists N: \forall n \geqslant N:\ \ \#\{\lambda \in \sigma(P_n^{-1}R_n):\ |\lambda - 1| \geqslant \epsilon\} \leqslant N;$$

that is, the CG iterations applied to the system (3) *converge to the solution at a superlinear rate.*

Proof. Let us study the eigenvalues of the matrix $\Delta_n = R_n - P_n$, which is related to $P_n^{-1}R_n$ by the formula $P_n^{-1}R_n = I_n + P_n^{-1}\Delta_n$. If we set $g(z) = f(z^{-1})$ in the assumptions of Theorem 3.3, we get the splitting

$$R_n - T_n(|f|^2) = E_n' + E_n'' \tag{12}$$

with $\|E_n''\|_2 < \delta$ arbitrary and

$$E_n' = \begin{pmatrix} * & & \\ & 0 & \\ & & * \end{pmatrix},$$

whose nonzero blocks have size N_1 depending only on δ.

Since P_n is the correction in the class $\mathcal{T}$ of $T_n(|f|^2)$, it can also be chosen N_2 such that for $n \geqslant N_2$

$$T_n(|f|^2) - P_n = \tilde{E}_n' + \tilde{E}_n'', \tag{13}$$

where $\|\tilde{E}_n''\|_2 < \delta$ and $\tilde{E}_n'$ vanishes except in the $N_2 \times N_2$ leading and trailing principal submatrices (compare [3]). From Eqs. (12) and (13) it follows

$$\Delta_n = E_n' + E_n'' + \tilde{E}_n' + \tilde{E}_n'',$$

$$\mathrm{rk}(E_n' + \tilde{E}_n') \leqslant N = 2\ \max(N_1, N_2),\ \|E_n'' + \tilde{E}_n''\|_2 < 2\delta.$$

Since f is continuous and has no root on the unit circle, the minimum value of $|f|^2$ for $|z| = 1$ must be positive; this yields the uniform invertibility of P_n, that is

$$\|P_n^{-1}\|_2 \leqslant c \quad \forall n,$$

for a suitable constant c (see [3]).

By the Courant–Fischer minimax characterization and the Cauchy interlacing theorem [20], it can be deduced that all the eigenvalues of $P_n^{-1}\Delta_n$ lie in the interval $(-\epsilon, \epsilon)$ with $\epsilon = 2c\delta$, except N possible outliers. This also proves the clustering of $\sigma(P_n^{-1}R_n)$ around unity; it then suffices to follow the proof in [13] in order to deduce the superlinear convergence rate of the conjugate gradient method. □

By following the terminology used in [32], the statement claims that $\sigma(P_n^{-1}R_n)$ has a "proper cluster" at 1.

It turns out that the statement of Theorem 3.4 is still valid if we replace P_n by whatever approximation $\tilde{P}_n$ in the class $\mathcal{T}$ giving an error of order ϵ with respect to the 2-norm. This can be particularly useful if the generating function f

is not explicitly known, so that the computation of $|f|^2$ (and the construction of P_n) cannot be carried out exactly.

If $\{t_k\}_{k\in\mathbb{Z}}$ are the coefficients of the matrix $T_n(f)$, then the Laurent expansion of $|f|^2$ is given by

$$|f(z)|^2 = \sum_{j=-\infty}^{+\infty} a_j z^j, \quad a_j = \sum_{k=-\infty}^{+\infty} t_k t_{k+j},$$

we could approximate the desired values $\{a_j\}_{j=0,\dots,n-1}$ with the truncated sums

$$\tilde{a}_j = \sum_{k=-M}^{M} t_k t_{k+j},$$

with M appropriately chosen. The $\tilde{a}_j$'s will give the "Toeplitz part" of our preconditioner $\tilde{P}_n \in \mathcal{T}$: the approximation error $\|P_n - \tilde{P}_n\|_2$ can be made arbitrarily small by a suitable choice of the truncation level M.

In fact, since P_n and $\tilde{P}_n$ both belong to the same algebra we may consider

$$\|P_n - \tilde{P}_n\|_2 = \max_j |\lambda_j(P_n) - \lambda_j(\tilde{P}_n)|,$$

and the eigenvalues of P_n or $\tilde{P}_n$ are the discrete sine transform of the values $\{a_j\}$ and $\{\tilde{a}_j\}$, respectively [3]. Thus, the error affecting $\tilde{a}_j$ controls the 2-norm distance between P_n and $\tilde{P}_n$.

Finally, observe that all the results of this section also hold if $f(z)$ belongs to the Dini–Lipschitz class. It suffices to invoke Weierstrass' theorem in order to find two Laurent polynomials $a(z)$ and $b(z)$ satisfying the inequalities (9) at the beginning of the proof in Theorem 3.3, and then to notice that Eq. (13) is valid as well [29].

4. Numerical results

In this section we want to test the effectiveness of our preconditioning strategy by some preliminary MATLAB experiment.

It is worth pointing out that we cannot expect our method to be competitive with respect to other techniques which iteratively solve a nonsymmetric system without using the normal equations.

For example, the circulant preconditioner proposed in [24] directly approximates the nonsymmetric matrix $T_n(f)$; it can so be used in connection to the CGS method [30], which often outperforms other methods like CGN. We meet the only exception where $T_n(f)$ is a banded matrix with f vanishing on some point of the unit circle: our Theorem 3.2 still holds, while the convergence results of [24] do not apply.

Hence, the following experiments are mainly concerned with least-squares Toeplitz problems, by proving that the solution can be computed in a lower

number of iterations if the normal equations are preconditioned by a sine transform instead of a circulant approximation. In particular, the last numerical example deals to a regularized inverse problem related to image restoration, where the solution of the normal equations is required.

In the first three examples, we consider the overdetermined linear system $C_{m,n}\mathbf{x} = \mathbf{y}$ whose right-hand side is $\mathbf{y} = (1, \ldots, 1)^{\mathrm{T}}$, for increasing values of n and $m = 2n$: we often set n as a decremented power of two, since in this case the $(n+1)$-point sine transform underlying the inversion in the $\mathscr{T}$ class becomes faster. The starting vector is the null vector, and the iterations are stopped when $\|C_{m,n}^{\mathrm{T}}(\mathbf{y} - C_{m,n}\mathbf{x})\|_2 < 10^{-12}$. We apply the CGN method with $\mathscr{T}$ or circulant preconditioning; if $C_{m,n}$ is partitioned as

$$\begin{pmatrix} C_n^{(1)} \\ C_n^{(2)} \end{pmatrix},$$

the circulant preconditioner is defined as

$$\begin{pmatrix} P_n^{(1)} \\ P_n^{(2)} \end{pmatrix},$$

each $P_n^{(i)}$ being the minimizer of $\|P_n - C_n\|_F$, as firstly proposed by Chan [14].

Example 1. $C_{m,n}$ is banded, well-conditioned and generated by

$$f(z) = -z^3 + 2z^2 + 9z + 3 - 2z^{-1} - 3z^{-2} + z^{-3}.$$

We sketch in Table 1 below the number of iterations performed by the various preconditioning methods.

The conclusions of Theorem 3.2 give $4\max(d^+, d^-) - 2 = 10$: our experimental results are close enough to such estimate, explaining why the sine transform works better.

Example 2. $C_{m,n} = T_{m,n}(f)$, where $f(z)$ is the rational function

$$\frac{1 + 0.7z}{1 - 0.9z} + \frac{1 - 0.8z^{-1}}{1 + 0.7z^{-1}},$$

the subdiagonal entries of C_n decay like 0.9^k, so that for small values of n the

Table 1

n	$\mathscr{T}$	Circulant
31	11	17
63	11	17
127	11	17
255	11	16

approximation of $T_{m,n}(f)^{\mathrm{T}}T_{m,n}(f)$ by $T_n(|f|^2)$ is not accurate. This explains the poor performance of our method for $n = 31$ (Table 2).

Example 3. The expression of the kth diagonal c_k of $C_{m,n}$ is

$$c_k = \begin{cases} 1/k^2 & \text{for } k \geqslant 1, \\ 2 & \text{for } k = 0, \\ 1/k^3 & \text{for } k \leqslant -1, \end{cases}$$

$C_{m,n} = T_{m,n}(f)$ where f is a nonrational Wiener function: our approach keeps being superior with respect to circulant preconditioning (Table 3).

Example 4. $C_n = T_n(f)$ where $f(z) = (1-z)^2(2-z^{-1})(3+z^{-1})$; now $\mathbf{y}$ is the right hand side corresponding to the true solution $\mathbf{x} = (1,\ldots,1)^T$, and the circulant preconditioner is just the optimal approximation to C_n. Our method becomes the most efficient for square nonsymmetric systems too, since Theorem 3.2 still applies while the assumption $|f(z)| > 0$ (required by the circulant preconditioning) is not verified for every z in the unit circle. In this situation it makes sense to compare the $\mathscr{T}$ preconditioned CGN method with the CGS method, which in the other examples is the best choice for a square linear system (Table 4).

Actually, the number of iterations increases with n due to the ill-conditioning of C_n.

Example 5. In this example the elements of $C_n = T_n(f)$ are

Table 2

n	$\mathscr{T}$	Circulant
31	18	13
63	9	13
127	6	13
255	5	12

Table 3

n	$\mathscr{T}$	Circulant
31	10	15
63	8	13
127	8	12
255	8	11

Table 4

n	$\mathscr{T}$	Circulant	
		CGS	CGN
31	9	13	19
63	11	17	25
127	13	21	35
255	16	28	51

$$c_k = \frac{\sigma}{\sqrt{\pi}} e^{-\sigma^2 k^2} \text{ if } |k| \leqslant 8, \quad c_k = 0 \text{ otherwise,}$$

where $\sigma = 40/303$.

Setting $n = 100$, we consider the vector $\mathbf{x}$ representing the image plotted in Fig. 1; define $\mathbf{y} = C_n\mathbf{x} + \boldsymbol{\eta}$, where $\boldsymbol{\eta}$ is a normally distributed noise vector whose relative norm is 10^{-3}. A plot of $\mathbf{y}$ is presented in Fig. 2.

The matrix C_n is very ill-conditioned, so that reconstructing $\mathbf{x}$ from $\mathbf{y}$ becomes an inverse problem; as an example, see in Fig. 3 the solution directly computed from $\mathbf{y}$ by a circulant preconditioned CG method. It is therefore

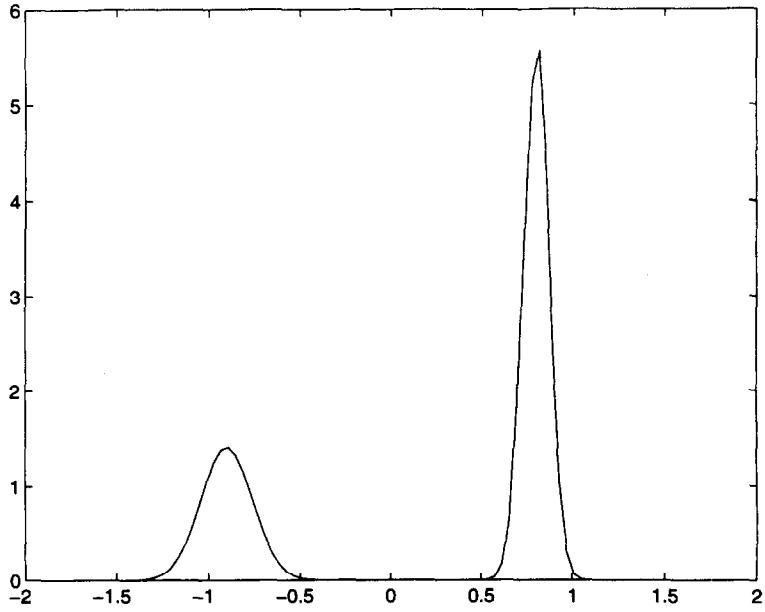

Fig. 1.

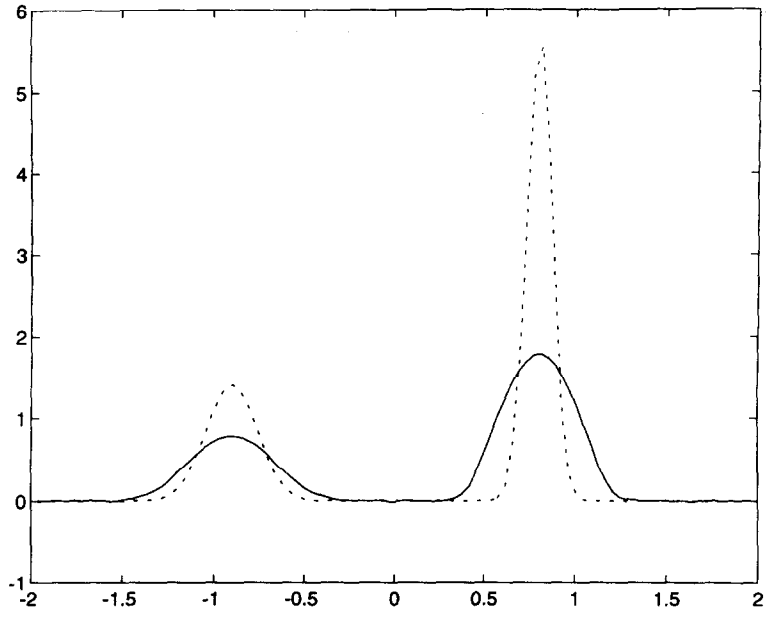
6
5
4
3
2
1
0
-1
-2
-1.5
-1
-0.5
0
0.5
1
1.5
2

Fig. 2.

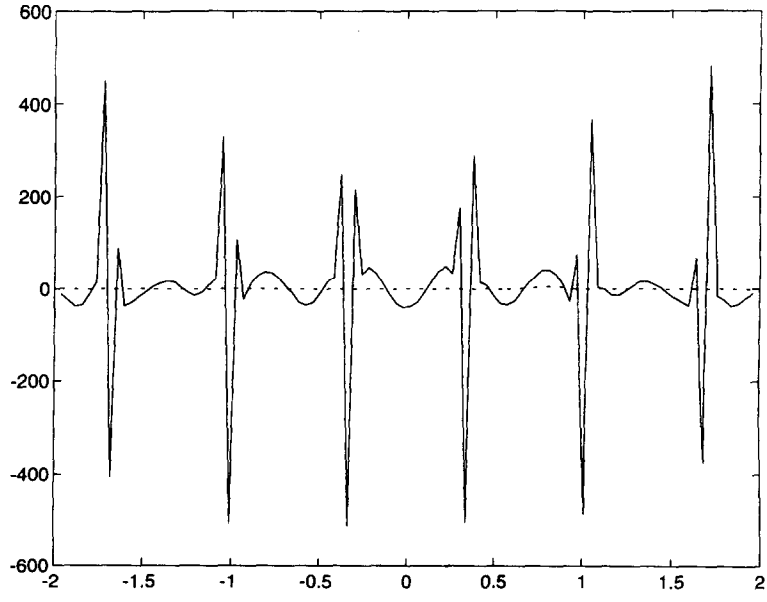
600
400
200
0
-200
-400
-600
-2
-1.5
-1
-0.5
0
0.5
1
1.5
2

Fig. 3.

necessary to apply a regularization technique; precisely, we search for the solution of the normal equations $H^{\mathrm{T}}H\tilde{\mathbf{x}} = H^{\mathrm{T}}\mathbf{y}$ where

$$H = \begin{pmatrix} C_n \\ \mu L \end{pmatrix},$$

μ being the regularization parameter (0.1 in this case) and $L = T_n(2 - z - z^{-1})$ discretizing the second-derivative operator, in order to make the approximate solution $\tilde{\mathbf{x}}$ smooth enough.

The CGN method can be applied in connection to the preconditioners

$$P_n = \tau_n(|f|^2) + \mu^2 L^2$$

belonging to $\mathscr{T}$, and

$$P_n = \mathrm{circ}(C_n)^2 + \mu^2 \mathrm{circ}(L)^2$$

where $\mathrm{circ}(\cdot)$ is the optimal circulant approximating matrix, as proposed in [8].

The iterations were stopped when the relative residual of the normal equations became less than 10^{-3}; in Figs. 4 and 5 we report the solutions computed by the two methods, after 4 and 8 iterations, respectively.

As it can be seen, the same precision is obtained by the $\mathscr{T}$ preconditioner in fewer iterations.

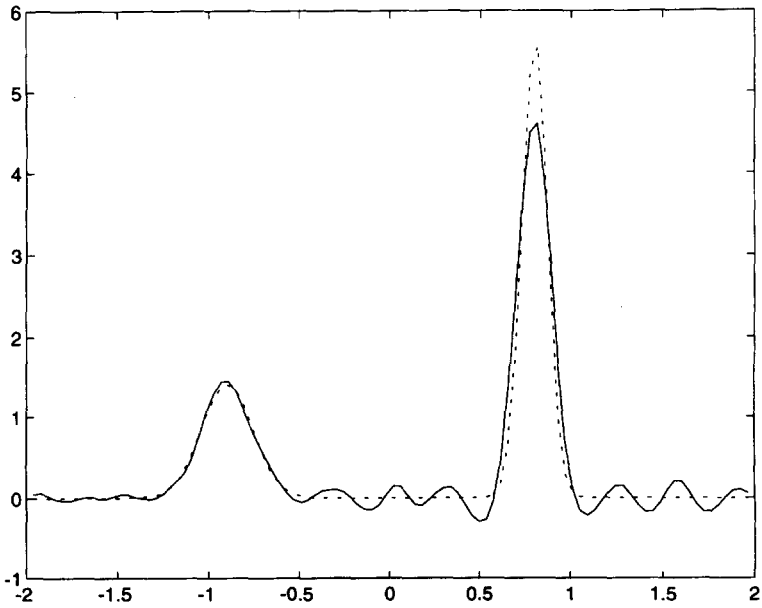

Fig. 4.

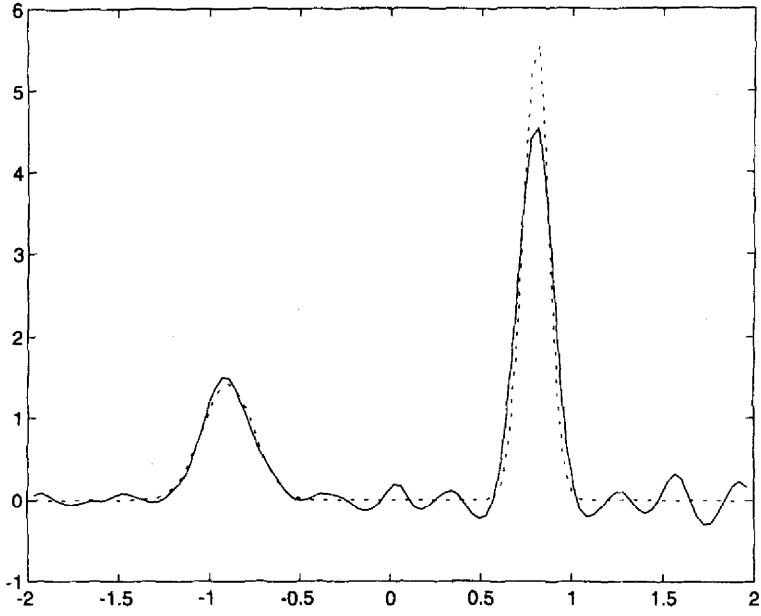

Fig. 5.

5. Generalization to block Toeplitz matrices

In some applications (multichannel signal processing, two-dimensional image restoration, PDE's) the elements defining each diagonal of a Toeplitz matrix are no longer scalars, but general $p \times p$ matrices: this is the classical case of *block Toeplitz matrices* like

$$C_{m,n} = \begin{pmatrix} c_0 & \cdots & c_{-n+1} \\ \vdots & \ddots & \vdots \\ c_{m-n} & & c_0 \\ \vdots & \ddots & \vdots \\ c_{m-1} & \cdots & c_{m-n} \end{pmatrix} \quad \text{with } c_j = (c_{r,s}^{(j)})_{r,s=1}^{p}. \tag{14}$$

For such matrices too, one can naturally define a generating function in terms of its diagonals: the main complication consists of the range of the function. In fact, its values don't belong to the complex scalar field, but to the algebra of complex $p \times p$ matrices. Precisely, $C_{m,n}$ is associated to the function $f : S^1 \to \mathbb{C}^{p\times p}$ if

$$\sum_{j=-\infty}^{+\infty} \|c_j\|_2 < +\infty \tag{15}$$

and the equality $f(z) = \sum_{j=-\infty}^{+\infty} z^j c_j$ is componentwise fulfilled. If so, we still use the notation $C_{m,n} = T_{m,n}(f)$.

A sensible simplification arises when the blocks defining the diagonals of a block Toeplitz matrix have in turn a strong structure: this is the case of *doubly Toeplitz matrices* [22,25] characterized by the pattern

$$C_{m,n} = \begin{pmatrix} c_0 & \cdots & c_{-n+1} \\ \vdots & \ddots & \vdots \\ c_{m-n} & & c_0 \\ \vdots & \ddots & \vdots \\ c_{m-1} & \cdots & c_{m-n} \end{pmatrix}, \quad c_j = \begin{pmatrix} c_{j,0} & c_{j,-1} & \cdots & c_{j,-p+1} \\ c_{j,1} & c_{j,0} & \ddots & \vdots \\ \vdots & \ddots & \ddots & c_{j,-1} \\ c_{j,p-1} & \cdots & c_{j,1} & c_{j,0} \end{pmatrix}.$$

If we consider the generating function of such a matrix, we obtain a function $f : S^1 \to \mathbb{C}^{p\times p}$ whose range is contained in the subspace of $p \times p$ Toeplitz matrices. That is, for every $z \in S^1$ the matrix $f(z)$ is Toeplitz itself and, of course, we could associate to it a scalar generating function $\phi(z,\cdot) : S^1 \to \mathbb{C}$, depending on z.

It is more practical to view ϕ as a bivariate function going from $S^1 \times S^1$ to $\mathbb{C}$ and to say that $C_{m,n}$ is directly generated by $\phi(z,w)$, by writing $C_{m,n} = T_{m,n,p}[\phi]$ (or simply $T_{n,p}[\phi]$ if $m = n$).

Clearly, $C_{m,n}$ is completely determined by the two-dimensional sequence $\{c_{j,k}\}$, whose entries represent the Fourier coefficients of ϕ:

$$c_{j,k} = \frac{1}{4\pi^2} \int\int_{\mathcal{Q}} \phi(e^{ix}, e^{iy}) e^{-i(jx+ky)}\, dx\, dy,$$

$\mathcal{Q}$ being the square $[-\pi,\pi]^2$ in $\mathbb{R}^2$. In this case, we still refer to ϕ as a *Wiener function.*

5.1. Structure properties

We point out that most of the results presented in the previous section about the structure of the product of Toeplitz matrices are directly extendable to the block case. Moreover, if the Toeplitz pattern is present in the inner blocks too, such results can be re-stated in a particular form in order to exploit the additional information.

For example, the proof of Theorem 3.1 does not care about the nature of the elements of $T_{n,m}(a)$ and $T_{m,n}(b)$, it still holds if such entries belong to a generic noncommutative algebra. It suffices to replace the ordinary transpose by the block transpose

$$C^*_{m,n} = \begin{pmatrix} c_0 & \cdots & c_{m-n} & \cdots & c_{m-1} \\ \vdots & \ddots & & \ddots & \vdots \\ c_{-n+1} & \cdots & c_0 & \cdots & c_{m-n} \end{pmatrix}$$

if $C_{m,n}$ is given by Eq. (14).

The new statement becomes

Theorem 5.1. *If $a(z)$ and $b(z)$ are $p \times p$ matrix Laurent polynomials such that*

$$a(z) = a_{-N}z^{-N} + \cdots + a_0 + \cdots + a_M z^M$$
$$b(z) = b_{-M}z^{-M} + \cdots + b_0 + \cdots + b_N z^N,$$

then $T_n(a \cdot b)$ and $T_{n,m}(a) \cdot T_{m,n}(b)$ differ by a $\mathscr{B}_{M,N,p}$ matrix.

In the doubly Toeplitz case, we obtain the following version.

Corollary 5.1. *Let $\alpha(z,w)$ and $\beta(z,w)$ be scalar functions defined as*

$$\alpha(z,w) = \sum_{\substack{-N \leqslant j \leqslant M \\ -L \leqslant k \leqslant K}} a_{j,k} z^j w^k, \qquad \beta(z,w) = \sum_{\substack{-M \leqslant j \leqslant N \\ -K \leqslant k \leqslant L}} b_{j,k} z^j w^k;$$

then the difference $T_{n,p}[\alpha \cdot \beta] - T_{n,m,p}[\alpha] \cdot T_{m,n,p}[\beta]$ can be written as $E_1 + E_2$, where $E_1 \in \mathscr{B}_{M,N,p}$ and $E_2 \in \mathscr{B}^{\Pi}_{K,L,n}$.

Proof. For all $z \in S^1$, let $a(z), b(z)$ be the $p \times p$ Toeplitz matrices

$$a(z) = T_p(\alpha(z,\cdot)), \quad b(z) = T_p(\beta(z,\cdot)),$$

these are matrix polynomials satisfying the assumptions of Theorem 5.1. It follows

$$T_n(a \cdot b) - T_{n,m,p}[\alpha] \cdot T_{m,n,p}[\beta] = E_1, \tag{16}$$

with $E_1 \in \mathscr{B}_{M,N,p}$, since $T_{n,m}(a) = T_{n,m,p}[\alpha]$ and $T_{m,n}(b) = T_{m,n,p}[\beta]$. Moreover, Theorem 3.1 can be applied for each z to the product $a(z) \cdot b(z)$

$$T_p(\alpha(z,\cdot) \cdot \beta(z,\cdot)) - a(z) \cdot b(z) = e(z), \tag{17}$$

where $e(z) \in \mathscr{B}_{K,L,1}$. Taking the two members of Eq. (17) as generating functions of suitable block Toeplitz matrices yields

$$T_{n,p}[\alpha \cdot \beta] - T_n(a \cdot b) = T_n(e)$$

with $T_n(e) \in \mathscr{B}^{\Pi}_{K,L,n}$ by construction. By recalling Eq. (16) we get the thesis

$$T_{n,p}[\alpha \cdot \beta] - T_{n,m,p}[\alpha] \cdot T_{m,n,p}[\beta] = E_1 + E_2,$$

where $E_2 = T_n(e)$. □

A slight complication arises when we go to examine full matrices. In order to reduce our study to the banded case, we need a generalization of Lemma 3.1

allowing us to estimate the norm of a block Toeplitz matrix in terms of the $p \times p$ matrix function that generates it. Assume that $r : S^1 \to \mathbb{C}^{p\times p}$ satisfies Eq. (15); define the quantity

$$|||r||| := \max_{z\in S^1} \|r(z)\|_2, \tag{18}$$

it is not difficult to check that $|||\cdot|||$ is a sub-multiplicative norm on matrix functions. We are now ready to prove the extension of Lemma 3.1.

Lemma 5.1. *Let $r(z)$ be a matrix-valued function defined on the unit circle satisfying the condition* (15). *If $|||r|||$ is defined as in Eq.* (18), *then the inequality*

$$\|T_{m,n}(r)\|_2 \leqslant |||r|||$$

holds.

Proof. Since $T_{m,n}(r)$ is generated by $r(z)$, we have the following integral representation for the block r_j of $T_{m,n}(r)$

$$r_j = \frac{1}{2\pi}\int_{-\pi}^{\pi} r(\mathrm{e}^{\mathrm{i}\theta})\mathrm{e}^{-\mathrm{i}j\theta}\,\mathrm{d}\theta, \quad j = -n+1,\ldots,m-1, \tag{19}$$

where the equality is satisfied componentwise.

In the following, assume that every vector of np elements be partitioned according to the block pattern of $T_{m,n}(r)$

$$\mathbf{u} = (\mathbf{u}^{(0)}\ldots\mathbf{u}^{(n-1)})^*, \quad \mathbf{u}^{(k)} = (u_{k,0}\ldots u_{k,p-1})^{\mathrm{T}}.$$

We may derive from Eq. (19) an integral expression for the element $v_{j,h}$ of the vector $T_{m,n}(r)\mathbf{u}$:

$$\begin{aligned}
v_{j,h} &= \sum_{k,l}[r_{j-k}]_{h,l}u_{k,l} = \sum_{k,l}\frac{1}{2\pi}\int_{-\pi}^{\pi}[r(\mathrm{e}^{\mathrm{i}\theta})]_{h,l}\mathrm{e}^{\ \mathrm{i}(j-k)\theta}u_{k,l}\,\mathrm{d}\theta \\
&= \sum_{k=0}^{n-1}\frac{1}{2\pi}\int_{-\pi}^{\pi}\mathrm{e}^{-\mathrm{i}(j-k)\theta}\sum_{l=0}^{p-1}[r(\mathrm{e}^{\mathrm{i}\theta})]_{h,l}u_{k,l}\,\mathrm{d}\theta \\
&= \sum_{k=0}^{n-1}\frac{1}{2\pi}\int_{-\pi}^{\pi}\mathrm{e}^{-\mathrm{i}(j-k)\theta}[r(\mathrm{e}^{\mathrm{i}\theta})\mathbf{u}^{(k)}]_h\,\mathrm{d}\theta \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi}\mathrm{e}^{-\mathrm{i}j\theta}[r(\mathrm{e}^{\mathrm{i}\theta})\sum_{k=0}^{n-1}\mathbf{u}^{(k)}\mathrm{e}^{\mathrm{i}k\theta}]_h\,\mathrm{d}\theta \\
&= \frac{1}{\sqrt{2\pi}}\langle[r(\mathrm{e}^{\mathrm{i}\theta})\mathbf{u}(\mathrm{e}^{\mathrm{i}\theta})]_h, \mathrm{e}^{(j)}\rangle;
\end{aligned}$$

$\langle\cdot,\cdot\rangle$ is the hermitian scalar product in $L^2(S^1)$, $\mathbf{u}(e^{i\theta})$ is the vector trigonometric polynomial $\sum_{k=0}^{n-1}\mathbf{u}^{(k)}e^{ik\theta}$, $e^{(j)}$ is the jth orthonormal function $(1/\sqrt{2\pi})e^{ij\theta}$.

The squared norm of the vector $T_{m,n}(r)\mathbf{u}$ can then be computed as

$$\sum_{j,h}|v_{j,h}|^2 = \frac{1}{2\pi}\sum_{h=0}^{p-1}\sum_{j=0}^{m-1}\left|\langle[r(e^{i\theta})\mathbf{u}(e^{i\theta})]_h, e^{(j)}\rangle\right|^2$$

$$\leqslant \frac{1}{2\pi}\sum_{h=0}^{p-1}\left\|[r(e^{i\theta})\mathbf{u}(e^{i\theta})]_h\right\|_{L^2}^2, \tag{20}$$

having applied the Bessel inequality to the expansion of $[r(e^{i\theta})\mathbf{u}(e^{i\theta})]_h$ with respect to the trigonometric basis.

Moreover,

$$\frac{1}{2\pi}\sum_{h=0}^{p-1}\left\|[r(e^{i\theta})\mathbf{u}(e^{i\theta})]_h\right\|_{L^2}^2 = \frac{1}{2\pi}\sum_{h=0}^{p-1}\int_{-\pi}^{\pi}\left|[r(e^{i\theta})\mathbf{u}(e^{i\theta})]_h\right|^2 d\theta$$

$$= \frac{1}{2\pi}\int_{-\pi}^{\pi}\sum_{h=0}^{p-1}\left|[r(e^{i\theta})\mathbf{u}(e^{i\theta})]_h\right|^2 d\theta$$

$$= \frac{1}{2\pi}\int_{-\pi}^{\pi}\|r(e^{i\theta})\mathbf{u}(e^{i\theta})\|_2^2\, d\theta$$

$$\leqslant \frac{1}{2\pi}|||r|||^2\int_{-\pi}^{\pi}\|\mathbf{u}(e^{i\theta})\|_2^2\, d\theta, \tag{21}$$

by the definition (18).

The last integral can be manipulated as

$$\int_{-\pi}^{\pi}\sum_{l=0}^{p-1}\left|\sum_{k=0}^{n-1}u_{k,l}e^{ik\theta}\right|^2 d\theta = 2\pi\sum_{l=0}^{p-1}\int_{-\pi}^{\pi}\left|\sum_{k=0}^{n-1}u_{k,l}e^{(k)}\right|^2 d\theta = 2\pi\sum_{l=0}^{p-1}\left\|\sum_{k=0}^{n-1}u_{k,l}e^{(k)}\right\|_{L^2}^2,$$

by Pythagoras' theorem, the inner L^2-norm is equal to $\sum_{k=0}^{n-1}|u_{k,l}|^2$. Recalling Eqs. (20) and (21) we deduce

$$\|T_{m,n}(r)\mathbf{u}\|_2^2 \leqslant \frac{1}{2\pi}|||r|||^2\cdot 2\pi\sum_{l=0}^{p-1}\sum_{k=0}^{n-1}|u_{k,l}|^2 = |||r|||^2\cdot\|\mathbf{u}\|_2^2$$

for every vector $\mathbf{u}$, which proves the thesis. □

If $f, g : S^1 \to \mathbb{C}^{p\times p}$ are arbitrary functions in the Wiener class, it is now straightforward to parallel the proof of Theorem 3.3 in order to get the following generalization.

Theorem 5.2. *For every $\epsilon > 0$ there exist N, M, ν such that for $n \geqslant \nu$*

$$T_{n,m}(g) \cdot T_{m,n}(f) - T_n(g \cdot f) = E' + E'',$$

where $\|E''\|_2 < \epsilon$ and $E' \in \mathscr{B}_{M,N,p}$.

Corollary 5.2. *If $\gamma(z,w)$ and $\phi(z,w)$ are Wiener functions, then for every $\epsilon > 0$ there exist N, M, K, L, ν such that for $n, p \geqslant \nu$*

$$T_{n,m,p}[\gamma] \cdot T_{m,n,p}[\phi] - T_{n,p}[\gamma \cdot \phi] = E_1 + E_2 + E_3,$$

with $E_1 \in \mathscr{B}_{M,N,p}$, $E_2 \in \mathscr{B}^{\Pi}_{K,L,n}$ and $\|E_3\|_2 < \epsilon$.

Proof. Define $g(z) = T_p(\gamma(z,\cdot))$, $f(z) = T_p(\phi(z,\cdot))$: they satisfy Eq. (15). We can apply Theorem 5.2 to $T_n(g \cdot f)$ obtaining

$$T_{n,m,p}[\gamma] \cdot T_{m,n,p}[\phi] - T_n(g \cdot f) = E' + E'' \tag{22}$$

with $E' \in \mathscr{B}_{M,N,p}$ and $\|E''\|_2 < \epsilon/2$. Moreover, the same theorem gives

$$g(z) \cdot f(z) = T_p(\gamma(z,\cdot) \cdot \phi(z,\cdot)) + e_1(z) + e_2(z)$$

with $e_1(z) \in \mathscr{B}_{K,L,1}$ and $\|e_2(z)\|_2 < \epsilon/2$ for all z, whence $|||e_2||| < \epsilon/2$ by definition (18). Thus

$$T_n(g \cdot f) = T_{n,p}[\gamma \cdot \phi] + T_n(e_1) + T_n(e_2) \tag{23}$$

where $T_n(e_1) \in \mathscr{B}^{\Pi}_{K,L,n}$ and $\|T_n(e_2)\|_2 \leqslant |||e_2||| < \epsilon/2$ by Lemma 5.1. By Eqs. (22) and (23) we deduce the desired relation, by setting $E_1 = E'$, $E_2 = T_n(e_1)$ and $E_3 = E'' + T_n(e_2)$. □

5.2. The block preconditioner

The preconditioning in the $\mathscr{T}$ class remains effective in the two-level setting. The basic idea still consists of finding a suitable approximation to the "normal equations" matrix $C^{\mathrm{T}}_{m,n} C_{m,n}$. Recall from Section 2 that our preconditioner $P_n \in \mathscr{T}_{n,p}$ must have a quadrantally symmetric pattern, which is equivalent to the condition

$$P_n = \tau_{n,p}[\theta] \quad \text{with } \theta(z,w) = \theta(z^{-1},w) = \theta(z,w^{-1}), \tag{24}$$

for a suitable Wiener function θ. Let us study the structure of $C^{\mathrm{T}}_{m,n} C_{m,n}$ by using the results of the previous section

$$C^{\mathrm{T}}_{m,n} C_{m,n} = T_{n,m,p}[\tilde{\phi}] \cdot T_{m,n,p}[\phi], \tag{25}$$

where $\phi(z,w)$ is the bivariate generating function of the doubly Toeplitz matrix $C_{m,n}$ and $\tilde{\phi}(z,w) = \phi(z^{-1}, w^{-1})$, as it can be easily checked, generates the transpose.

By Corollary 5.2, the right-hand side of Eq. (25) can be written as

$$T_{n,p}[\tilde{\phi} \cdot \phi] + E_1 + E_2 + E_3,$$

where E_1 and E_2 are sparse and the norm of E_3 can be taken arbitrarily small (or zero in the banded case). Thus, the natural definition of P_n should be the following

$$P_n = \tau_{n,p}[\tilde{\phi} \cdot \phi],$$

it is then evident that P_n is well defined if the condition (24) is verified for $\theta(z,w) = \phi(z,w)\phi(z^{-1}, w^{-1})$. For this, it suffices to assume that

$$\forall j: \quad c_j = c_j^{\mathrm{T}}, \tag{26}$$

i.e. the matrix $C_{m,n}$ is symmetric at the inner level; this implies $\phi(z,w) = \phi(z,w^{-1})$ whence the condition (24) for θ. The hypothesis (26) is not too restrictive, being very common in practice.

Moreover, the positive definiteness of P_n is ensured by the nonnegativity of its generating function, according to the properties recalled in Section 2: in fact,

$$\theta(z,w) = \phi(z,w)\phi(z^{-1}, w^{-1}) = \phi(z,w)\overline{\phi(z,w)} = |\phi(z,w)|^2$$

for all z, w on the unit circle. In particular, P_n and $T_{n,p}[\theta]$ are both well-conditioned if ϕ does not vanish on any point of $S^1 \times S^1$.

We can start to analyze the banded case: let $R_n = C_{m,n}^{\mathrm{T}} C_{m,n}$ and assume that ϕ be a bivariate polynomial.

Theorem 5.3. *Let* $\phi(z,w) = \sum_{\substack{-M \leqslant j \leqslant N \\ -K \leqslant k \leqslant L}} c_{j,k} z^j w^k$; *then*

$$P_n^{-1} R_n = I_{np} + E_n$$

with $\mathrm{rk}(E_n) \leqslant (M+N)p + (K+L)n$.

Proof. Define $\Delta_n = R_n - P_n$ and observe that

$$\Delta_n = (R_n - T_{n,p}[\theta]) + (T_{n,p}[\theta] - \tau_{n,p}[\theta]);$$

the two terms both belong to $\mathscr{B}_{M,N,p} + \mathscr{B}^{\Pi}_{K,L,n}$, by Corollary 5.1 and ([17], Theorem 3.2), respectively. Recalling that $P_n^{-1} R_n = I_{np} + \Delta_n$, the thesis follows. □

Corollary 5.3. *The CG method applied to the normal equations* (3) *with preconditioner* P_n *converges in exact arithmetic within* $O(n+p)$ *iterations.*

Finally, when P_n is well-conditioned we have the clustering of the spectrum of $P_n^{-1}R_n$ around the unity: the proof is based on Corollary 5.2 and ([17], Theorem 3.3).

Theorem 5.4. *Let $\phi(z,w)$ be a Wiener function without zeroes on $S^1 \times S^1$. If Σ is the spectrum of $P_n^{-1}R_n$, then*

$$\forall \epsilon > 0 \ \exists \nu: \forall n,p \geqslant \nu \quad \#\{\lambda \in \Sigma: |\lambda - 1| \geqslant \epsilon\} \leqslant O(n+p).$$

According to the notions introduced in [32], the statement above implies the existence of a "general (or weak) cluster" at 1 for Σ. Notice the difference with respect to the scalar case, where the cluster is "proper": in the block setting we cannot prove a superlinear convergence result for the preconditioned CGN method.

Example 6. We report from [16] a block version of the experiment performed in Example 5. Here $C_{m,n} = T_{m,n,p}[\phi] \in \mathbb{R}^{4096\times 4096}$, with

$$\phi(z,w) = \sum_{|j|,|k| \leqslant 8} e^{-0.1(j^2+k^2)} z^j w^k \quad \text{and} \quad m = n = p = 64.$$

Now $\mathbf{x}$ represents the 2-d image of 64×64 pixels given in Fig. 6, $\boldsymbol{\eta}$ is chosen as in Example 5 in order to get the blurred and noisy image $\mathbf{y} = C_{m,n}\mathbf{x} + \boldsymbol{\eta}$ plotted in Fig. 7.

Define

$$H = \begin{pmatrix} C_{m,n} \\ \mu I \end{pmatrix},$$

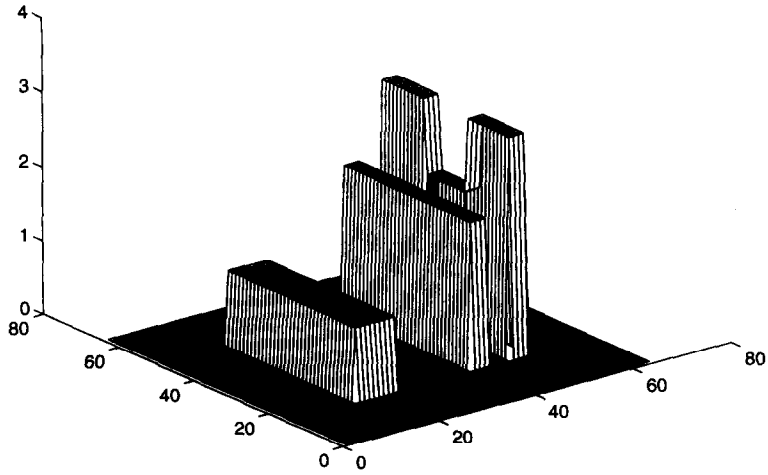

Fig. 6.

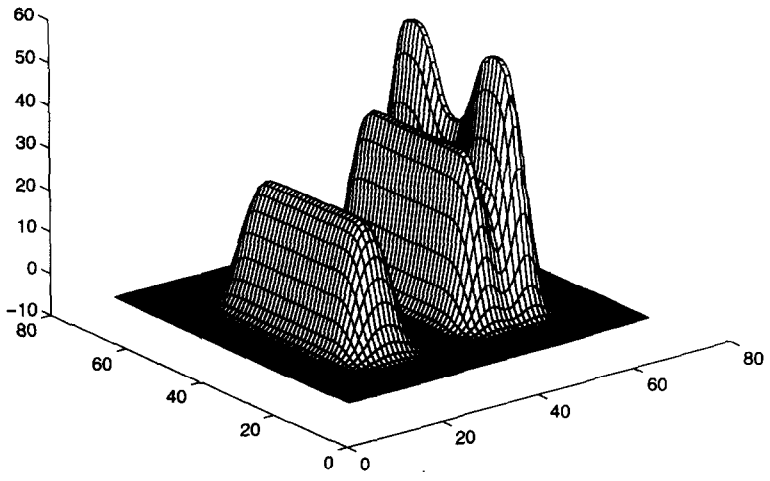

Fig. 7.

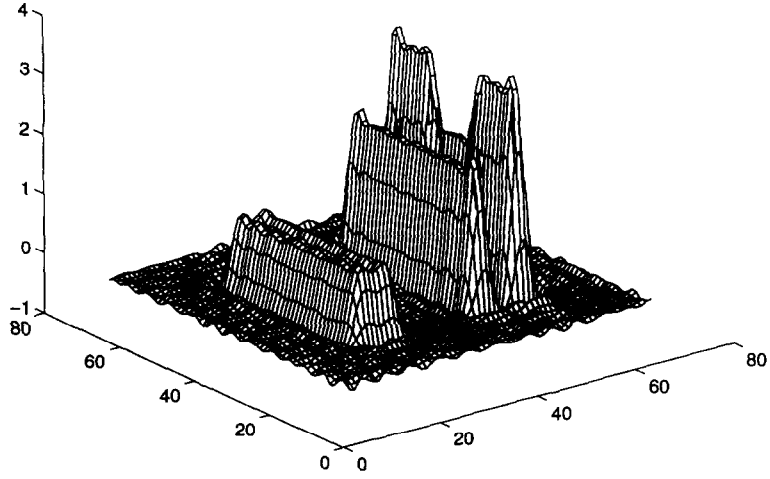

Fig. 8.

where $\mu = 0.1$ is the regularization parameter and I is the identity matrix (in this way we control the size of the solution). If we solve the normal equations $H^{\mathrm{T}}H\tilde{\mathbf{x}} = H^{\mathrm{T}}\mathbf{y}$ by the CGN method with the preconditioner

$$P_n = \tau_{n,p}[|\phi|^2 + \mu^2],$$

after 13 iterations we obtain the reconstruction shown in Fig. 8 and a residual error reduced by a factor 10^{-3}.

References

[1] G. Ammar, W. Gragg, Superfast solution of real positive definite Toeplitz systems, SIAM J. Matrix Anal. Appl. 9 (1988) 61–76.
[2] D. Bini, M. Capovani, Spectral and computational properties of band symmetric Toeplitz matrices, Linear Algebra Appl. 52 (1983) 99–126.
[3] D. Bini, F. Di Benedetto, A new preconditioner for the parallel solution of positive definite Toeplitz systems, Proceedings of the Second Annual SPAA, ACM Press, Crete, Greece, 1990, pp. 220–223.
[4] E. Boman, I. Koltracht, Fast transform based preconditioners for Toeplitz equations, SIAM J. Matrix Anal. Appl. 16 (1995) 628–645.
[5] J.R. Bunch, Stability of methods for solving Toeplitz systems of equations, SIAM J. Sci. Stat. Comp. 6 (1985) 349–364.
[6] R.H. Chan, X.Q. Jin, A family of block preconditioners for block systems, SIAM J. Sci. Stat. Comp. 13 (1992) 1218–1235.
[7] R.H. Chan, J.G. Nagy, R.J. Plemmons, FFT-based preconditioners for Toeplitz-block least squares problems, SIAM J. Numer. Anal. 30 (1993) 1740–1768.
[8] R.H. Chan, J.G. Nagy, R.J. Plemmons, Circulant preconditioned Toeplitz least squares iterations, SIAM J. Matrix Anal. Appl. 15 (1994) 80–97.
[9] R.H. Chan, J.G. Nagy, R.J. Plemmons, Displacement preconditioner for Toeplitz least squares iterations, Elec. Trans. Numer. Anal. 2 (1994) 44–56.
[10] R.H. Chan, K.P. Ng, Toeplitz preconditioners for Hermitian Toeplitz systems, Linear Algebra Appl. 190 (1993) 181–208.
[11] R.H. Chan, M. Ng, Conjugate gradient methods for Toeplitz systems, SIAM Rev. 38 (1996) 427–482.
[12] R.H. Chan, M. Ng, C.K. Wong, Sine transform based preconditioners for symmetric Toeplitz systems, Linear Algebra Appl. 232 (1996) 237–260.
[13] R.H. Chan, G. Strang, Toeplitz equations by conjugate gradients with circulant preconditioner, SIAM J. Sci. Stat. Comp. 10 (1989) 104–119.
[14] T. Chan, An optimal circulant preconditioner for Toeplitz systems, SIAM J. Sci. Stat. Comp. 9 (1988) 766–771.
[15] F. Di Benedetto, Analysis of preconditioning techniques for ill-conditioned Toeplitz matrices, SIAM J. Sci. Comp. 16 (1995) 682–697.
[16] F. Di Benedetto, Iterative solution of Toeplitz systems by preconditioning with the discrete sine transform, in: F. Luk (Ed.), Proceedings of the SPIE Conference on Advanced Signal Processing Algorithms, Architectures, and Implementations, San Diego, CA, vol. 2563, 1995, pp. 302–312.
[17] F. Di Benedetto, Preconditioning of block Toeplitz matrices by sine transforms, SIAM J. Sci. Comp. 18 (1997) 499–515.
[18] F. Di Benedetto, G. Fiorentino, S. Serra, C.G. preconditioning for Toeplitz matrices, Computers Math. Appl. 25 (1993) 35–45.
[19] F. Di Benedetto, S. Serra Capizzano, A unifying approach to abstract matrix algebra preconditioning, Numer. Math. (to appear).
[20] G.H. Golub, C. Van Loan, Matrix Computations, The Johns Hopkins University Press, Baltimore, MD, 1983.
[21] M.R. Hestenes, E. Stiefel, Methods of conjugate gradients for solving linear systems, J. Res. Nat. Bur. Standards 49 (1952) 409–435.
[22] J.H. Justice, A Levinson-type algorithm for two-dimensional Wiener filtering using bivariate Szegö polynomials, Proc. IEEE 65 (1977) 882–886.
[23] T.K. Ku, C.C.J. Kuo, On the spectrum of a family of preconditioned block Toeplitz matrices, SIAM J. Sci. Stat. Comp. 13 (1992) 948–966.

[24] T.K. Ku, C.C.J. Kuo, Spectral properties of preconditioned rational Toeplitz matrices: The nonsymmetric case, SIAM J. Matrix Anal. Appl. 14 (1993) 521–544.
[25] B.C. Levy, M.B. Adams, A.S. Willsky, Solution and linear estimation of 2-D nearest-neighbor models, Proc. IEEE 78 (1990) 627–641.
[26] L. Mertens, H. Van de Vel, A special class of structured matrices constructed with the Kronecker product and its use for difference equations, Linear Algebra Appl. 106 (1988) 117–147.
[27] J.G. Nagy, R.J. Plemmons, Iterative image restoration using FFT-based preconditioners, Proceedings of The Allerton Conference, University of Illinois, Champaign, Urbana, 1992.
[28] S.V. Parter, On the distribution of the singular values of Toeplitz matrices, Linear Algebra Appl. 80 (1986) 115–130.
[29] S. Serra, Superlinear PCG methods for symmetric Toeplitz systems, Math. Comp. (to appear).
[30] P. Sonneveld, CGS: A fast Lanczos-type solver for nonsymmetric linear systems, SIAM J. Sci. Stat. Comp. 10 (1989) 115–130.
[31] E. Tyrtyshnikov, Influence of matrix operations on the distribution of eigenvalues and singular values of Toeplitz matrices, Linear Algebra Appl. 207 (1994) 225–249.
[32] E. Tyrtyshnikov, A unifying approach to some old and new theorems on distribution and clustering, Linear Algebra Appl. 232 (1996) 1–43.